DOCUMENT RESUME

ED 238 022 CS 208 066

AUTHOR Reed, W. Michael; Burton, John K.
TITLE Effective and Ineffective Evaluation of Essays: Perceptions of College Freshmen.
PUB DATE [81]
NOTE 23p.
PUB TYPE Reports - Research/Technical (143) -- Guides - Classroom Use - Guides (For Teachers) (052)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Essays; *Evaluation Methods; Higher Education; Informal Assessment; Questionnaires; Student Attitudes; *Student Reaction; *Writing Evaluation; *Writing Instruction; *Writing Research
IDENTIFIERS Freshman Composition

ABSTRACT

To assess the effectiveness of various composition evaluation methods, a study asked two freshman English classes to respond to three questionnaires on the writing of personal essays and methods of evaluation they had experienced. Responses indicated that all students had a fear of having their essays evaluated. Evaluation methods considered a disadvantage by a majority of the students were the circling of errors, marking a paper only for what is wrong or what is right with it, and using terminology not understood by students. Practices considered effective included peer evaluations, student-instructor conferences, and the occasional use of ungraded writing. (Appended are the student responses to the questionnaires.) (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED238022

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Effective and Ineffective Evaluation of Essays: Perceptions of College Freshmen

by

W. Michael Reed and John K. Burton

Division of Curriculum and Instruction
College of Education
Virginia Tech

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
W. Michael Reed

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

W. Michael Reed
400 War Memorial Gym
Virginia Tech
Blacksburg, VA 24061
(703) 961-5587/5428

CS 208066

A major assumption underlying studies of the composing process during the last several decades has been that writing performance is primarily a function of information stored in permanent memory (long-term memory). That is, if writers have in their permanent memory writing-related information, such as events to write about, writing skills to communicate with, and strategies by which to generate content and formalize that content, they should be successful at completing a writing task.

This assumption is a basic underpinning of composing process models, beginning with Rohman's in 1965.[1] His model describes the sequential writing behaviors, interpretations of which have been modified by Emig,[2] Perl,[3] Sommers,[4] and Flower and Hayes.[5] Although distinctions regarding stage-engagement (prewriting, writing, and rewriting) of good writers versus poor writers have been made, these models have not taken into account the effect that evaluation has on the composing process. And even though intuitively we know that willingness and motivation (potential effects of evaluation) must affect our approach to and completion of writing tasks, our models have reflected only processes, not factors affecting these processes. How does willingness, for example, affect writers searching their permanent memory, as depicted by Flower and Hayes?[6] Because our models do not address this issue, a logical source to find out about willingness is our students.

We need to ask our students what kind of evaluation helps them the most, or, better yet, what kind of evaluation offends them the most or helps them the least? Because we need to view essay feedback as having a cyclical effect, we should turn to our students and ask them what kind of evaluation is the most effective? After all, if our feedback is not effective to them specifically, our well-meaning attempts at providing effective feedback are lost.

Although evaluation has been a major part of our composition courses, we have not always considered the effect that evaluation can have. Studies dealing with willingness to write have been somewhat sparse, but we can turn to general feedback studies to answer our questions on the effects that evaluation can have on students. And because writing essays involves people writing about their feelings, their ideas, and their beliefs, it is more likely that a debilitative effect will occur than when students are doing other types of work. A D, for example, on an essay probably has a deeper effect than a D on a multiple-choice history midterm because the writers can easily interpret an essay grade as judgment of their feelings, beliefs, and ideas rather than merely not knowing answers. Because of the complexities involved both in writing an essay and giving an adequate, fair, and effective evaluation of that essay, we need, first, to ask our students what kinds of evaluation work best for them and, second, to provide them evaluative "moments" to maximize the time and energy they, the writers, and we, the evaluators, invest in a writing assignment.

Studies indicate that consequences (grades) can become very powerful in determining people's success at a given task.[7] For example, a negative evaluation can result in a conditioned effect so that the student will avoid the pain (in this case, yet another D on an essay), by not doing the task at all (I won't write another essay!). In this case, writers are more concerned with their self-image than with the consequences (Even though I received a zero for that assignment, it was because I didn't do it, rather than doing it and receiving a low grade). And even if these writers decide to attempt a writing task, their willingness (motivation) might be so low that it could very easily become a deterrent to completing the task.[8] Simply put, past success or failure at a specific type of task affects willingness and performance at similar tasks in the future.[9]

Certainly the problem we face is providing our writers with effective feedback that not only describes the quality of their writing (accompanied with directions on how to go about improving their writing) but also does not negatively affect their willingness to improve that particular essay or subsequent essays. Effective feedback provides the writer with a combination of what is right and what is wrong with the essay.[10] It has been found that writers whose essays were thoroughly evaluated with comments improve their writing moreso than those writers whose essays are sparsely marked.[11] Comments on personal essays should be detailed (give examples of the examples you expect in an essay), factual (avoid remarks that might reflect a simple

difference of opinion), positive (encourage the writers with what they have done right), and clear (use terminology that the writer understands).[12]

If, indeed, we view evaluation as having a cyclical effect--that is, the evaluation of one essay should lead the writer to improving the next essay--we need to insure that evaluation is having the effect we intend it to have. One way to determine the effect of evaluation is to ask our students what seems to help them revise their current essays and what seems to help them write the next one. After all, because we view learning as cumulative, we should also view evaluation as cumulative.

Procedures. Two intact English classes of first-quarter average freshman writers at a southeastern university were recently asked to respond to three questionnaires (open-ended and forced choice) related to the writing of personal essays and past and present methods of evaluation they have experienced. The first questionnaire (see Table 1) was administered at the end of the third week (n = 47) before any essays had been turned in; the second questionnaire (see Table 2), at the end of the fifth week (n = 42) after one essay had been evaluated; and the third questionnaire (see Table 3), at the end of the ninth week (n = 44) after three essays had been evaluated.

(Insert Table 1)

Evaluation Methods. Based on the first questionnaire, which focused on the types of evaluation their essays received during high

school, all writers expressed a fear of having their essays evaluated. We often interpret a certain degree of fear as reflecting the degree of challenge a task offers.[13] However, 83 percent specifically attributed this fear to some form of evaluation. Past evaluation fell into one of three categories: essays were graded for grammar only (45 percent), for content only (17 percent), or for both content and grammar (38 percent). The freshmen whose essays had been evaluated for grammar only went on to state that they felt this kind of evaluation was limited because it did not give them credit for the ideas they produced for that essay. A common reaction was, what I had to say was not important, but rather whether I got my commas right or whether the subject and verb agreed. Those students whose essays had been evaluated for content only, on the other hand, felt they were not given credit for correct grammar and usage. They further felt that their ideas were being judged. If we combine these concerns, a logical remedy might be to grade for both. In fact, at the end of the quarter, 88 percent of these freshmen felt that essays should be evaluated for both content and grammar (see Table 3).

(Insert Table 2)

Lack of Instruction. Because most people feel that writing involves complex processes and requires practice, we need to provide as many writing opportunities as possible. Simply put, if we are holding our student writers accountable for what we expect their

writing to encompass, they need numerous writing occasions and instruction. Of these 47 writers, only 19 percent had had courses designed specifically for the learning of writing compositions.

The Essence of Evaluation. It is not so much that personal essays are being evaluated that creates much of this fear, but rather how these essays are being evaluated. Students achieve less when they receive more feedback to incorrect answers in the form of criticism.[14] Again, this reaction relates to the nature of essay writing: if writers include their opinions, beliefs, and ideas and the content of their essays is criticized, then it is not surprising that they feel they are being judged rather than their papers. Feedback should be designed so that, by being told what is right and what is wrong, writers will make adjustments for later writing tasks. No feedback should attack their ideas, beliefs, or opinions, but rather just their treatment (both logical and grammatical) of these personal, textual inclusions.

(Insert Table 3)

Dead-end Evaluation. There are many "short-hand" approaches to providing feedback, usually benefitting the evaluator more than the writer. These writers were asked to react to several forms of evaluation by indicating whether they felt each was an advantage or disadvantage and why. Circling, considered a disadvantage by 75 percent of the surveyed writers, generally signifies that something is wrong. It may be faulty punctuation, no punctuation at all, a

misspelled word, incorrect usage, inconsistent verb tense, almost anything! Although circling is effective in letting the writers know something is wrong (that circle, often in red, stands out among all those words), it does a poor job letting them know why that "branded" spot is wrong. If a circle indicates several possible types of mistakes, writers would have to have a good foundation in order to figure it out. If such a foundation is needed, then they probably would not have committed the error in the first place! If writers do not have this foundation, there is a good chance they will not be able to understand the mistake and, most likely, will not attempt to correct it. Both groups of people lose: the evaluators, because their attempts at providing feedback fall on "deaf ears," and the writers, because they do not understand enough to effectively apply the feedback to improving their writing.

Because marking a paper for what is wrong with it is perhaps the least time-consuming way to evaluate a paper, we probably use this type of feedback more than other types. This approach perhaps reflects both our sense of duty to let our students know what is wrong and our assumption that, once knowing what is wrong, the students will make the effort necessary for correcting their mistakes. Seventy-seven percent of the surveyed students felt that this type of evaluation was a disadvantage. The underlying reasons for this opinion were that it, first, depressed them and, second, it did not acknowledge what they did right. Fifty-nine percent of the writers, on the other hand, felt

that marking a paper for what is right with it was a disadvantage because they also wanted to know what was wrong. An obvious solution to turning these disadvantages into advantages is to provide feedback that contains both methods of evaluation: what is right and what is wrong with the essay. By incorporating both, we are providing much more feedback than if we were using only one of the two. Intuitively, we know that the more information we receive about the quality of our work (areas we are having problems with and areas that we have command of), the more deliberate and directed our future efforts at improving our writing will be.

Only 20 percent of the surveyed freshmen felt that giving directions on how to correct errors was a disadvantage. Because 80 percent felt this was an advantage, we perhaps need to incorporate these directions in our evaluations. For those of us who feel this might give too much away and prevent them from a problem-solving exercise, we need to remind ourselves that they have been continually engaged in problem-solving exercises throughout the writing of that essay and many of these exercises have been successful, as the correct portions of their essays will indicate. For the 20 percent who felt this method was a disadvantage because it restricts their individual approaches to correcting the problem, we should tell our students that our directions are merely suggestions and they can correct the errors in their own way.

When we use terminology as part of our feedback, we are potentially facing a major problem: what happens if they do not understand the term? Certainly when we use such terms, we are assuming they understand what the term means. Although we might feel we are justified in expecting them to know what the term means, there is always the possibility that they will not. Again, our efforts might be ill spent. Eighty percent of the surveyed writers felt this method was a disadvantage mainly because the term might be unfamiliar to the student. The term in the questionnaire was misplaced modifier. If the writers are not adequately familiar with this term (what do I need to do to make it not misplaced? why is it a modifier? what is a modifier?), their attempts at correcting it might confound the error. Other similar terms include parallelism, subject-verb agreement, diction, awk, stitled, trite!

Providing Effective Feedback. We first need to become aware of the effect that our evaluative efforts are having on our writers. If we sense that some of our efforts are having little or no effect, we need to review, modify, and test out our revised methods. Thus far, we have been concerned with methods of formal evaluation. Another awareness we need to have is that feedback does not always have to be formal, and, in fact, some informal types of feedback can provide the immediacy that formal methods cannot provide.

One type of informal feedback, for example, is peer evaluations of essays. According to the surveyed students, 89 percent felt that peer

evaluations were helpful (see Table 3). Usually peer evaluations--generally in groups of two or three--take place several days before an essay is actually due. There are several advantages to incorporating peer evaluations in writing classes: first, they allow writers the opportunity to have their essays examined for errors without being penalized for these mistakes; second, there is sufficient time for them to correct the problems prior to turning them in for a grade; and third, peer evaluations allow the writers to see that other writers are having the same problems they are having (the mistake is not just "in the eyes of the evaluator") and that their colleagues are overcoming similar problems (If he can overcome the problem, then so can I!).

Another form of effective feedback is the student-instructor conference. To help insure that a formal evaluation of an essay does not fall on "deaf ears," we can incorporate in our writing classes this type of conference which usually involves the student and the instructor going over a graded essay. Ninety-eight percent of these freshmen felt that the student-instructor conferences were beneficial. Besides receiving individual attention, the writers were also able to ask the instructor to clarify comments or to provide instruction on a particular area--two things that are especially difficult when evaluation is limited to merely handing back a graded essay.

Another type of student-instructor conference is actually grading an essay during the session. This method has several effects: first,

it demythologizes the grading of an essay for the writer; second, it clarifies once again to the writers what the criteria are for the essays; and, third, it provides a one-to-one interaction between the student and instructor. For these 30-minute grading conferences, the surveyed writers, prior to the conference, had filled out a questionnaire which addressed the time they spent on the essay, what they perceived to be weaknesses in their essays as well as the strengths, the grade they felt their essays deserved, and a justification for these grades (see Table 4). The writers first read their essays outloud (this allowed the instructor to deal with the clarity and logic of the content). Afterwards, the writer and the instructor discussed the strengths and weaknesses of the essay. The instructor then read the essay for grammatical accuracies, and then a grade was assigned. The instructor explained to the writer why the essay deserved that grade.

(Insert Table 4)

The grades the writers had given their essays were not looked at until after a grade had been assigned. It was found that of the 45 writers, the grades the writers had given their essays were very close to the grades the instructor had given the essays. In fact, 96 percent of the pairs of grades (the student's grade and the instructor's grade) were within two levels of each other, when the difference, for example, between a B+ and a B- represents two grade levels (see Table 5 for grade comparisons). The closeness between the grades indicates that

writers know how well they have done on an essay prior to turning it in, and, if this is this case, we need to incorporate more methods of evaluation other than simply handing back a graded essay. If we do not, our evaluations simply confirm what they already know about their writing. To insure that they go beyond what they already know, we need to give them directions and insights into how to write even better.

(Insert Table 5)

If we subscribe to the notion that writing is a complex activity that requires practice, then we need to consider ungraded writing. These surveyed freshmen wrote eight-page collections of writing every two weeks. These collections were not graded, but simply reacted to. Even though 62 percent of these writers still felt some fear when turning in these ungraded writings (see Table 3), 83 percent felt more comfortable than when turning in essays to be evaluated. Ungraded writing seems to satisfy two concerns: first, the need to provide writers with many opportunities to write; and, second, to minimize their fear of having their work evaluated. Another impression was that these writers often took more chances in their writing when they were not faced with a grade.

Ungraded writing, though, should be only one of many components of a writing class. When asked at the end of the quarter to choose between 1) only graded writing, 2) only ungraded writing, and 3) both graded and ungraded writing, 67 percent wanted the combination and 21

percent wanted only graded writing. Despite the degree of comfort the writers felt with ungraded writing, grades still prove to be an incentive to write as well as possible.

Summary. If we want our evaluative efforts to be as effective as possible, we need to find out from our students what types of evaluation work best for them when they attempt to apply feedback to improving their writing. Once we find out, we need to quickly incorporate these methods as part of our evaluation of their essays. Also, the purpose of feedback is to accurately respond to all that they have produced, in terms of both what is right and what is wrong. Although we might assume that whatever we do not mark up should be considered right, student writers will most often look only for reactions from the instructor to get a sense of the quality of the paper. If numerous reactions are negative, then this one-sided evaluation can have an effect that we do not intend to make. Also besides using effective formal evaluative methods, we also need to incorporate informal methods, such as peer evaluations (they can learn much from one another), student-instructor conferences, and ungraded writing exercises.

Notes

1. D. G. Rohman, "Prewriting: The Stage of Discovery in the Writing Process," College Composition and Communication, 16 (1965), pp. 106-112.

2. J. Emig, The Composing Processes of Twelfth Graders (Urbana, IL: NCTE, 1971).

3. S. Perl, "The Composing Processes of Unskilled Writers," Research in the Teaching of English, 13 (1979), pp. 317-336.

4. N. Sommers, "Revision Strategies of Student Writers and Experienced Adult Writers," College Composition and Communication, 31 (1980), pp. 378-388.

5. L. Flower and J. Hayes, "The Cognition of Discovery: Defining a Rhetorical Problem," College Composition and Communication, 31 (1980), pp. 21-32.

6. Flower and Hayes.

7. S. Lee, "Motivational Effects of Normative Feedback Following Success-Failure Experience on Attribute Identification Performance," American Educational Research Journal, 13 (1976), pp. 255-266.

8. M. E. Ururoglu and H. Walberg, "Motivation and Achievement: A Qualitative Synthesis," American Educational Research Journal, 16 (1979), pp. 375-389.

9. R. M. Latta, "Interactive Effects of Initial Achievement Orientation and Prior Success Feedback on the Mastery of Subsequent Difficult and Easy Tasks," American Educational Research Journal, 15 (1978), pp. 17-24.

10. C. Barringer and B. Gholson, "Effects of Type of Combination of Feedback upon Conceptual Learning by Children: Implications for Research in Academic Learning," Review of Educational Research, 49 (1979), pp. 459-478.

11. D. G. Kehl, "The Art of Writing Evaluative Comments on Student Themes," English Journal, 59 (1970), pp. 972-982.

12. C. Lynch and P. Klemans, "Evaluating Our Evaluations," College English, 40 (1978), pp. 166-180.

13. T. M. Wildman and J. K. Burton, "Integrating Learning Theory with Instructional Design, Journal of Instructional Development, 4 (1981), pp. 5-13.

14. J. Martin, D. Veldman, and L. M. Anderson, "Within-class Relationships Between Student Achievement and Teacher Behaviors," American Educational Research Journal, 17(1980), pp. 479-490.

Table 1. Freshman Responses to Experiences with Past Evaluations of Essays

n = 47

1. Do you have any fears about turning in personal essays to be evaluated?	100%--yes 0%--no
2. If you have any fears, to what do you attribute these fears?	17% (8)--to what other people told them about college freshman English 28% (13)--inconsistent grading in the past 45% (21)--expected harder grading in college 10% (5)--past performance
3. How were your essays evaluated in the past?	45% (21)--grammar only 38% (18)--both grammar and content 17% (8)--content only
4. Did you have courses designed specifically to instruct you in how to write compositions?	81% (38)--no 19% (9)--yes
5. Were you as anxious when turning in writing assignments for other subject-matter courses as when turning in essays for English class?	79% (37)--no 21% (10)--yes

Table 2. Freshman Responses to Certain Methods of Evaluation

n = 42

Explain the advantage or disadvantage of the following types of evaluation:

1. having something circled	75% (32) disadvantage	24--it did not explain what was wrong 8--too many had a depressing effect
	25% (10)--advantage	all--the error stands out
2. evaluating for only what is wrong with the essay	77% (33)--disadvantage	23--it has a depressing effect 10--it does not tell what is correct
	23% (9)--advantage	all--it lets the writer know what is wrong
3. evaluating for only what is right with the essay	59% (25)--disadvantage	8--it does not tell what is wrong with the essay 17--it gives the writer a false sense of total accomplishment
	41% (17)--advantage	all--it makes the student feel he/she knows something
4. giving directions on how to correct an error	20% (8)--disadvantage	all--these directions restricted the writer's individual ways to correct the error
	80% (34)--advantage	all--these directions let the student know how to go go about correcting the errors
5. using terms, such as misplaced modifier	80% (34)--disadvantage	all--such terms might be unfamilia
	20% (8)--advantage	all--such terms let the student know exactly what is wrong

Table 3. Freshman Responses to Current Methods

n = 44

1. Were peer evaluations of essays beneficial?	89% (39)--yes	26--others could help locate errors 3--students of same age could more closely identify with content 10--peers could help with content problems
	11% (5)--no	2--peers did not have sufficient background to help 1--peers held back criticism 1--peers wanted to change original ideas 1--peers were not interested in helping
2. Were the student-instructor conferences helpful?	98% (43)--yes	38--the instructor could further explain comments and give more direction 5--the instructor and the student got to know each other better
	2% (1)--no	1--the instructor and the student went over information the student already knew
3. Do you have any fears when turning in Bi-Weekly Writings?	62% (27)--yes 38% (17)--no	
4. Did you feel more comfortable or less comfortable when turning in Bi-Weekly Writings than when turning in personal essays?	88% (39)--more comfortable 12% (5)--less comfortable	
5. Do you feel grades are needed incentives to write better?	60% (26)--yes 40% (18)--no	
6. If you were to help design a writing program, which would you choose: only ungraded writing, only graded writing, a combination of both?	12% (5)--only ungraded 21% (9)--only graded 67% (30)--a combination	

Table 3. Freshman Responses to Current Methods

(continued)

n = 44

7. If you were to choose the method by which to evaluate essays, which would you choose: grammar only, content only, both grammar and content?	5% (2)--grammar only 7% (3)--content only 88% (39)--combination

Table 4. Questionnaire on Essay #3 Filled out by Student Prior to Student-Instructor Grading Conference*

---

1. How much time did you spend on this paper (including all three steps)?

2. After the peer evaluation of your rough draft, what did you do to improve your paper?

How successful were you?

3. What are the strengths of your paper?

4. What are the weaknesses, if any, of your paper?

5. What problems did you have when writing this paper?

What problems do you still have questions about?

6. What grade would you give your paper?

7. Explain why you would give your essay this grade. Base your answer on how well your essay meets the criteria established for essay grades.

Table 5. Students' Grades and Instructor's Grades on Essay #3

n = 45

(NOTE: 96% of pairs of grades were within two grade levels of each other--A = one grade level, A- = one grade level, B+ = one grade level, etc.)

Student	Student Grade	Instructor's Grade	Grade Level Difference*
1	B-	C+	1. 1-
2	D+	C-	2. 1+
3	B	A-	3. 2+
4	B**	C	4. 3-
5	C	C	5. 0
6	B	B-	6. 1-
7	C	C-	7. 1-
8	C	C+	8. 1+
9	D	D-	9. 1-
10	D+	D	10. 1-
11	B	B	11. 0
12	C	C-	12. 1-
13	A-**	B-	13. 3-
14	C	C+	14. 1+
15	B	B+	15. 1+
16	D+	C-	16. 1+
17	C	C	17. 0
18	B	B	18. 0
19	C+	C	19. 1-
20	C+	C-	20. 2-
21	B	B	21. 0
22	C	C+	22. 1+
23	B	B-	23. 1-
24	C-	D+	24. 1-
25	B	B	25. 0
26	B+	B	26. 1-
27	B	B-	27. 1-
28	C	C+	28. 1+
29	C	C-	29. 1-
30	C	C-	30. 1-
31	B-	B-	31. 0
32	B	B+	32. 1+
33	D+	D	33. 1-
34	C	C+	34. 1+
35	B	B-	35. 1-
36	B+	B+	36. 0
37	B	B-	37. 1-
38	A-	B	38. 2-
39	B-	C+	39. 1-
40	C	C	40. 0
41	C	C	41. 0
42	C	C-	42. 1-
43	D	D+	43. 1+

Table 5. Students' Grades and Instructor's Grades on Essay #3

(continued)

Student	Student Grade	Instructor's Grade	Grade Level Difference
44	C-	C	44. 1+
45	D+	C	45. 2+

*The "+" mark indicates that the instructor's grade was higher than the student's grade, and the "-" mark indicates that the instructor's grade was lower than the student's grade. The number before either a "+" or a "-" indicates the number of grade levels between the instructor's grade and the student's grade.

**The grade level difference between the instructor's grade and the student's grade is more than two grade levels.